Apprentissage et évolution

apprentissage :

- un joueur fait face à la même situation stratégique à plusieurs reprises et "apprend" le comportement optimal à force de jouer.
- plusieurs modèles différents, qui reposent sur différents mécanismes d'apprentissage :
  - imitation
    (faire ce qu'un autre joueur a fait si cela a donné un bon payoff)
  - empirisme :
    - le joueur anticipe le choix de son adversaire en se basant sur le comportement antérieur (passé) de l'adversaire.
    - le joueur choisit la stratégie qui lui a donné le plus de succès dans le passé.

exemple :

Chicken

2

1

	F	D
F	0,0	4,1
D	1,4	2,2

Jeu fictif ("Fictitious Play")

Modèle à 2 joueurs.

- Chaque joueur prédit la stratégie mixte qui sera utilisée par l'autre à la prochaine itération.
- Cette prédiction est basée uniquement sur la fréquence avec laquelle l'autre joueur a joué chacune de ses stratégies pures dans le passé.
- Le joueur choisit sa meilleure réponse à cette stratégie mixte.

		② F	② D	
①	F	0,0	4,1	p
	D	1,4	2,2	$1-p$
		q	$1-q$	

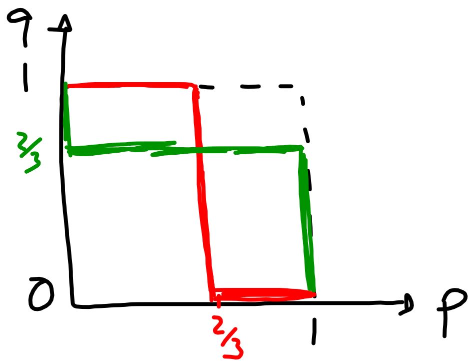
① $E\pi_1(F) = 0 \cdot q + 4(1-q) = 4 - 4q$
$E\pi_1(D) = 1 \cdot q + 2(1-q) = 2 - q$

choisit F lorsque $4-4q > 2-q$
$\rightarrow 2 > 3q$
$\rightarrow 2/3 > q$

choisit D lorsque $q > 2/3$

② $E\pi_2(F) = 4 - 4p$
$E\pi_2(D) = 2 - p$

choisit F lorsque $p < 2/3$
D lorsque $p > 2/3$



Si on simule le jeu fictif, le comportement des joueurs convergera-t-il vers un des EN ?
Si oui, lequel ?

Comme pour toute méthode itérative, il faut choisir un point de départ (condition initiale).

itération	①	p	②	q
1	F	$1/4$	F	$1/2$
2	D	$5/8$	F	$3/4$
3	D	$5/12$	F	$5/6$
4				
5				
6				

car $1/4 < 2/3$

car $1/2 < 2/3$

Le joueur 1, à la période $t+1$, anticipe que 2 jouera F avec probabilité q_{t+1} qui est un calcul de la fréquence avec laquelle 2 a joué F dans le passé :

$$q_{t+1} = \begin{cases} \dfrac{q_t \cdot t + 1}{t+1} & \text{si } ② \text{ a joué F à la dernière itération} \\[2ex] \dfrac{q_t \cdot t}{t+1} & \text{sinon} \end{cases}$$

Ainsi $q_2 = \dfrac{\frac{1}{2} \cdot 1 + 1}{1+1} = \dfrac{3}{4}$

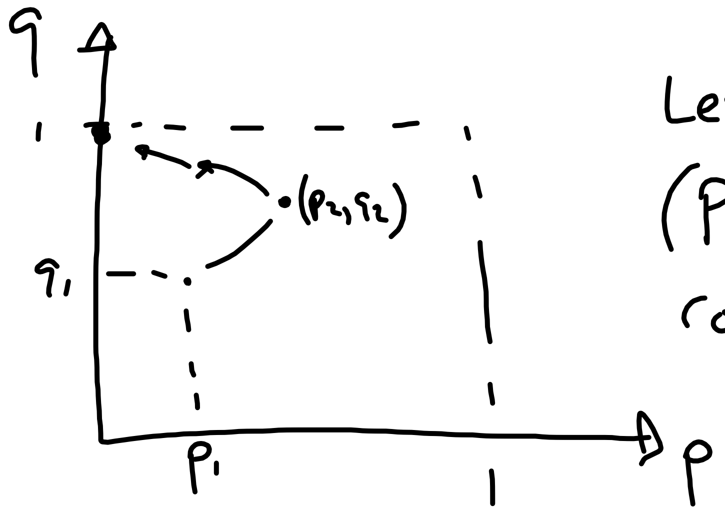
$$p_2 = \frac{\frac{1}{4} + 1}{1+1} = \frac{5}{8}$$

$$q_3 = \frac{2 \cdot \frac{3}{4} + 1}{3} = \frac{5}{6}$$

$$p_3 = \frac{\frac{5}{8} \cdot 2}{3} = \frac{5}{12}$$

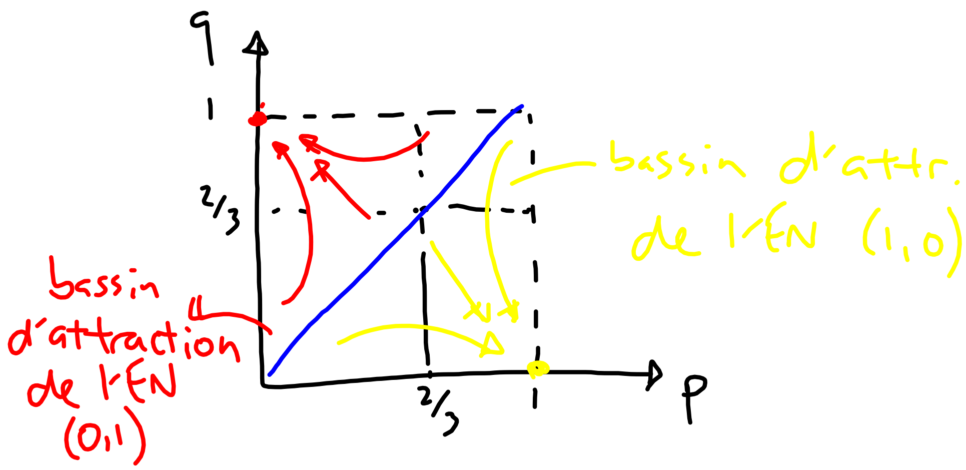
$$q_4 = \frac{5/6 \cdot 3 + 1}{3+1} = 7/8$$

$$p_4 = \frac{5/12 \cdot 3 + 0}{3+1} = 5/16$$



Les fréquences (p_t, q_t) convergent vers $(0,1)$

En répétant l'exercice avec des conditions initiales différentes, on serait en mesure de dire vers quel équilibre on se dirige pour chaque point de départ possible.



Les deux équilibres en stratégies pures sont stables (on peut y converger avec des conditions initiales appropriées).

L'équilibre $(2/3, 2/3)$ est instable (impossible d'y converger si on commence ailleurs qu'à $(2/3, 2/3)$).